**AP STAT- Ch. 3 -- 5 Quiz Review**

1. A survey of automobiles parked in the student and staff lots at a large university classified the brands by country of origin, as seen in the table below:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | Driver | |  |
|  |  | **Student** | **Staff** | TOTAL |
|  | **American** | 107 | 105 | 212 |
| Origin | **European** | 33 | 12 | 45 |
|  | **Asian** | 55 | 47 | 102 |
|  | TOTAL | 195 | 164 | 359 |

1. What is the marginal distribution of Origin? Make a bar graph.
2. What is the marginal distribution of Driver? Do not make a bar graph.
3. What percent of Students drove Asian cars?
4. What percent of Asian cars are driven by staff?
5. What percent of Staff drove Asian cars?
6. What percent of those surveyed were Students?
7. What percent of those surveyed drove American cars or were students?
8. What percent of those surveyed drive European cars and were staff?
9. What is the conditional distribution of Origin?
10. What is the conditional distribution of Driver?
11. Create a segmented bar chart for the conditional distribution of Driver.
12. Is there an association between Origin and Driver? Provide statistical evidence to support your claim.
13. Create a dotplot of the number of goals scored by each team in the first round of the California high school soccer playoffs. Then briefly describe the distribution.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 5 | 0 | 1 | 0 | 7 | 2 | 1 | 0 | 4 | 0 | 3 | 0 | 2 | 0 |
| 3 | 1 | 5 | 0 | 3 | 0 | 1 | 0 | 1 | 0 | 2 | 0 | 3 | 1 |

1. Create back-to-back stemplots of the following male and female heights. Compare & describe both distributions

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **MALE** | |  |  |  |  |  | **FEMALES** | | |  |  |  |
| 72 | 75 | 66 | 76 | 70 | 71 |  | 72 | 69 | 70 | 64 | 70 | 66 |
| 73 | 74 | 65 | 73 | 73 | 66 |  | 70 | 60 | 71 | 65 | 61 | 67 |
| 73 | 68 | 65 | 63 | 72 | 68 |  | 66 | 59 | 70 | 66 | 69 | 68 |
| 70 | 68 | 70 | 64 | 72 | 69 |  | 61 | 61 | 60 | 66 | 68 | 68 |
| 71 | 67 | 71 | 60 | 71 | 72 |  | 60 | 62 | 61 | 66 | 67 | 65 |

1. Find the 5# summaries and create parallel boxplots for the heights of males and females in question #2
2. Salaries of 2008 New York Yankees (in millions of dollars):

Rodriguez 28 Giambi 23.428

Jeter 21.6 Abreu 16

Petite 16 Rivera 15

Posada 13.1 Damon 13

Matsui 13 Mussina 11.071

Pavano 11 Farnsworth 5.917

Wang 4 Hawkins 3.75

Cano 3 Molina 1.875

Ensberg 1.75 Brackman 1.185

Betemit 1.165 Bruney 0.725

Traber 0.500 Cabrera 0.461

Hughes 0.406 Duncan 0.398

Henn 0.397 Kennedy 0.394

Karstens 0.393 Albaladejo 0.393

Ohlendorf 0.391 Chamberlain 0.390

Sanchez 0.390

1. Create a frequency histogram of the data above. Describe the distribution.
2. Based on this description, what measure of center and spread should you report?
3. Find the mean, standard deviation, 5# summary, and IQR
4. Create a cumulative frequency histogram.
5. Heights (in cm) of 58 randomly selected Canadian students who participated in a survey

166.5 170 178 163 150.5 169 171 166 190 183 178 161

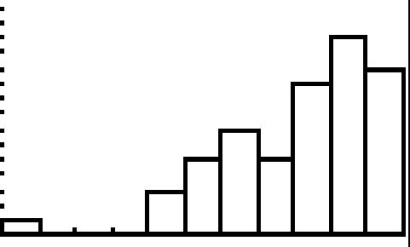
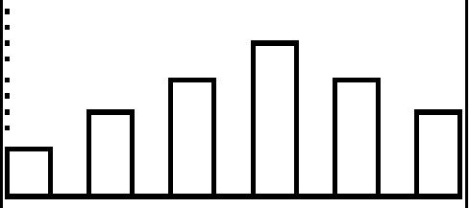
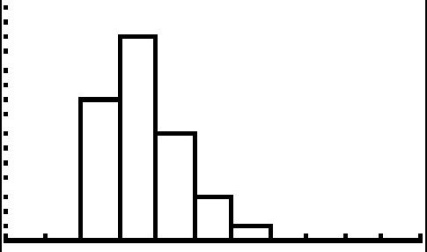
171 170 191 168.5 178.5 173 175 160.5 166 164 163 174

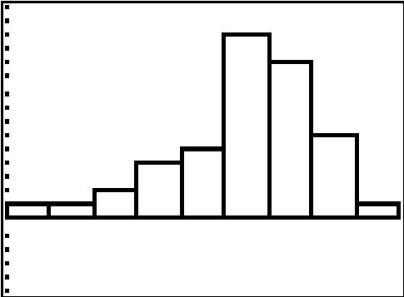
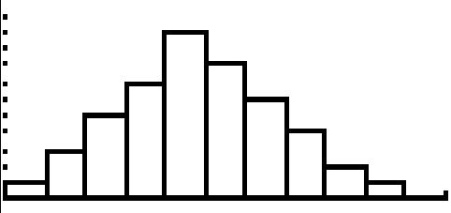
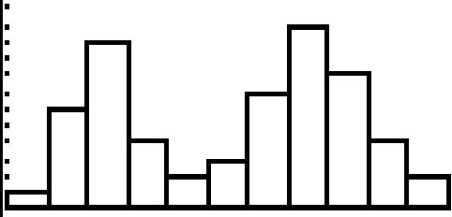
173 169 160 174 182 167 166 170 170 181 171.5 160

178 157 165 187 168 157.5 145.5 156 182 168.5 177 162.5

* 1. 185.5 151 159 177 171 176 177 181 186

1. Create a relative frequency histogram of the following data. Describe the distribution.
2. Based on this description, what measure of center and spread should you report?
3. Find the mean, standard deviation, 5# summary, and IQR
4. Use the following data. {30, 30, 30, 30, 30, 30, 30, 30}. Find the mean and standard deviation. Why is the standard deviation this value?
5. Describe the following distributions using the terms we learned in class. Scale on x-axis: (1, 12), bins = 1

1. Use the following data: {20, 23, 24, 27, 29, 31, 30, 33, 36, 37, 35, 40}
2. Calculate the following statistics:

Mean

Median

Range

IQR

Std. Dev.

1. Suppose we now add a new point to the data set: 60. Indicate whether adding the new point to the rest of the data made each of the summary statistics in part (a) increase, decrease, or stay about the same
2. A random sample of the heights of 24-34 year old women was taken (in inches). The following summary statistics were calculated.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Statistic** | **mean** | **st. dev.** | **min** | **Q1** | **med** | **Q3** | **max** |
| **Heights of 24-34 year old women** | **69.5** | **2.65** | **58** | **62** | **64** | **68** | **78** |

1. Based on the summary statistics would you describe the distribution as symmetric or skewed? Explain.
2. Are there any outliers present? Show all work.